Theor Appl Genet (2002) 105:440–448
DOI 10.1007/s00122-002-0945-x

J.C. Ho · S.R. McCouch · M.E. Smith

Improvement of hybrid yield by advanced backcross QTL analysis in elite maize

Received: 22 May 2001 / Accepted: 14 November 2001 / Published online: 19 June 2002

Abstract We applied an advanced backcross breeding strategy to identify quantitative trait loci (QTLs) of agronomic importance in a cross between two elite inbreds of maize, RD6502 (Mo17-type recurrent parent) and RD3013 (Iodent donor parent). Two hundred and four BC_2 families were scored at 106 SSR, 15 AFLP, and 38 *Heartbreaker* (MITE) loci. BC_2 testcrosses (TC) with B73 were phenotyped at six locations in the Midwest and N.Y. We detected four grain yield, six grain moisture, and three plant height QTLs at which the RD3013 allele had a favorable effect ($p < 0.05$). All four yield QTLs were selected as target introgressions in the development of BC_3TC families. As predicted by BC_2TC analysis, BC_3TC entries containing introgressions at *yld3.1* and *yld10.1* significantly outperformed non-carrier entries by 11.1% (15.6 bu/A at one location) and 6.7% (7.1 bu/A averaged across two locations), respectively, in replicated Midwestern trials ($p < 0.05$). Detection of *yld10.1* effects in the BC_2TC by spatial analysis (i.e., incomplete block, response surface, autoregressive, moving average or autoregressive moving average), but not by conventional single point analysis or interval mapping, indicated the utility of local environmental control for QTL mapping in unreplicated maize progeny. This work demonstrated that the advanced backcross QTL method can be applied to identify and manipulate useful QTLs in heterotic inbreds of elite maize. Genetic gains by this approach can be coupled with the maintenance and selection of favorable epistatic gene complexes by traditional hybrid breeding for maize improvement.

Communicated by H.C. Becker

J.C. Ho
Institute for Genomic Diversity, Cornell University,
157 Biotechnology Bldg, Ithaca, NY 14853-2703, USA

S.R. McCouch · M.E. Smith (✉)
Department of Plant Breeding, 252 Emerson Hall,
Cornell University, Ithaca, NY 14853-1901, USA
e-mail: mes25@cornell.edu
Tel.: +1-607-255-1654, Fax: +1-607-255-6683

Keywords Maize · Advanced backcross-QTL analysis · Grain yield · Spatial analysis · *Heartbreaker* miniature inverted transposable elements

Introduction

The advanced backcross (AB) method has been used for the simultaneous identification and transfer of favorable quantitative trait locus (QTL) alleles from unadapted material to elite lines (Tanksley and Nelson 1996). By this approach, QTLs having positive effects on key agronomic traits (e.g., fruit weight and soluble-solids content) were discovered in the wild species *Lycopersicon pimpinellifolium*, *Lycopersicon pennellii* and *Lycopersicon hirsutum*, and transferred to cultivated tomato (*Lycopersicon esculentum*) (de Vicente and Tanksley 1993; Eshed and Zamir 1994; Tanksley et al. 1996). Similarly, Xiao et al. (1998) and Moncada et al. (2001) generated interspecific BC_2 testcross and BC_2F_2 populations of rice (*Oryza sativa* L.), and determined that introgression of *Oryza rufipogon* QTL alleles into elite rice hybrid and inbred varieties enhanced grain yield by 17% and 14%, respectively.

In this study, we hypothesized that AB-QTL analysis can be extended to crosses between elite, heterotic inbreds of maize. In contrast to tomato and rice, maize is an allogamous crop species that encompasses abundant genetic variation (Darrah and Zuber 1986) despite a domestication bottleneck (Eyre-Walker et al. 1998) followed by intensive breeding. If high-order epistasis plays a major role in hybrid maize performance, marker-assisted selection (MAS) done in advanced backcross progeny may be superior to such selection in the F_2. In an F_2 population, recombination can dismantle the assembly of favorable epistatic gene combinations accumulated by traditional breeding (Allard 1996). Because an advanced backcross population is skewed toward the recurrent parent genome, favorable epistatic interactions among recurrent parent and tester alleles are less likely to be disrupted. Conversely, there is a higher probability

of detecting donor parent alleles with useful additive or dominant effects that will be expressed in a nearly isogenic background (Tanksley and Nelson 1996).

Commonly observed upon transfer of individual QTLs to isogenic backgrounds (Doebley et al. 1995; Long et al. 1995; Eshed and Zamir 1996; Laurie et al. 1997), epistasis has been established as a major mechanism of adaptation in various plant species (Allard 1996) and the genetic basis of heterosis in rice (Yu et al. 1997). Evidence of epistasis and its importance in maize has accumulated steadily as well (e.g., Lamkey et al. 1995; Allard 1996; Wolf and Hallauer 1997; Eta-Ndu and Openshaw 1999; Lukens and Doebley 1999). Through concurrent programs of parental line development and hybrid evaluation, maize breeders have selected and fixed favorable epistatic gene combinations in inbreds with excellent specific combining ability. The common practice of developing source populations by crossing inbreds from the same heterotic pattern (Melchinger et al. 1998) has served to maintain and accumulate favorable epistatic gene combinations, particularly those in linkage disequilibrium (Lamkey et al. 1995). For example, the exceptional performance of the widely grown hybrid B73 × Mo17 has been attributed in part to positive epistatic effects on heterosis (Wolf and Hallauer 1997). In another study, Stuber and Sisco (1991) identified yield QTLs in the F_2 progeny of two maize inbreds (Tx303 and Oh43), backcrossed the chromosomal segments into B73, and found that it was disadvantageous to transfer more than two or three segments to the elite line. The authors postulated that epistatic interactions among a larger number of introgressions can cause an unfavorable effect on yield.

In this experiment, we performed QTL analysis in a BC_2 population of elite maize for the improvement of hybrid yield. By crossing heterotic inbreds, we targeted QTLs that may not have been manipulated by conventional methods of line development based on crosses between closely related lines (Lamkey et al. 1995; Wolf and Hallauer 1997; Melchinger et al. 1998). By delaying selection until the BC_2 generation, we emphasized the detection of useful additive and dominant QTL effects (Tanksley and Nelson 1996) that otherwise might be confounded with novel and negative epistatic interactions in balanced populations (e.g., F_2, recombinant inbred lines). The primary objective of this study was to assess the efficiency of AB-QTL analysis for the identification and transfer of agronomically useful QTLs in heterotic gene pools of elite maize.

Materials and methods

Population development

To develop a BC_2 population, we selected two Cornell inbreds, RD6502 (Mo17-type) as a male recurrent parent and RD3013 (Iodent) as a female donor parent. RD3013 is an early maturing inbred that combines well with both RD6502 and B73, the tester used in this experiment. RD6502 was crossed with RD3013 to generate F_1 plants, which subsequently were backcrossed to RD6502 to generate 18 BC_1 families. Two or three plants from each BC_1 family were backcrossed again to RD6502 to generate a total of 39 BC_2 families. The following year, five or six plants from each BC_2 family were self-pollinated, backcrossed, and crossed to B73 to generate 204 sets of BC_2S_1, BC_3 and BC_2 testcross (TC) families, respectively. We applied minimal selection at the BC_1 and BC_2 stages by eliminating agronomically inferior plants (e.g., very late maturity or lack of coincidence between pollen shed and silk emergence) and advancing plants with good ear quality and seed set.

In 1998, the 204 BC_2TC families, one commercial check (Pioneer 3394), and three replicates of the original B73 × RD6502 hybrid were randomized and machine-planted in two-row plots (5.3×0.76 m^2 at 69,200 plants/ha) as unreplicated trials at six locations: Aurora, N.Y., Jerseyville, Ill, Kingston, N.Y., Pittsford, N.Y., Winterset, Iowa and York, Neb. Weed and insect control and nutrient management recommendations for each local area were followed. Five or six plants from each BC_3 family were self-pollinated in a summer nursery the same year.

Trait evaluation

At all locations, grain was combine-harvested and evaluated for *grain yield* (adjusted to 15.5% grain moisture), *grain moisture* (%) at harvest, and *yield-to-moisture* ratio. At N.Y. locations, plant height and standability data were collected after flowering on a plot basis. *Plant height* (cm) was based on the average of six competitive plants randomly sampled from each plot and measured from the soil surface to the tip of the central spike of each tassel. *Root lodging* was measured as the number of plants leaning more than 30 degrees from the vertical. *Stalk lodging* was measured as the number of plants broken at or below the primary ear node.

Molecular markers

To reconstitute BC_2 genotypes, the DNA of each BC_2S_1 family was extracted as a pooled sample from freshly harvested leaves of 20 plants according to Causse et al. (1994). Similarly, we obtained DNA of parental lines RD6502 and RD3013. The 204 BC_2 progeny were genotyped at 106 simple sequence repeat (SSR) loci (http://www.agron.missouri.edu) according to Causse et al. (1994). Gel electrophoresis and microsatellite detection by silver-staining were performed according to Panaud et al. (1996). SSR data were collected by visually scoring the presence or absence of the RD3013 allele at each marker locus.

In addition, we genotyped the population at 16 amplified fragment length polymorphism (AFLP) and 38 *Heartbreaker* (*Hbr*) miniature inverted transposable element (MITE) loci. For AFLP analysis, a two-step amplification strategy was performed according to Vuylsteke et al. (1999). After screening 14 *Pst*I/*Mse*I primer combinations for polymorphism between parental lines, we selected the four most informative primer pairs for mapping in the BC_2 population. For *Hbr* transposon display (TD), we followed the modified AFLP procedure described by Casa et al. (2000). After screening eight *Mse*I and *Bfa*I primer-enzyme combinations in the selective amplification step, we selected the four primer pairs that produced the highest number of polymorphic bands for mapping in the BC_2 population. All unlabeled primers

and adapters were synthesized by Research Genetics (http://www.resgen.com). The *Hbr*Int5 and *Pst*I selective primers were synthesized on an Applied Biosystem 392 (Applied Biosystem, Foster City, Calif.) and labeled with hexachloro-6-carboxyfluorescein (HEX) or 6-carboxyfluorescein (6-FAM) by the Cornell BioResource Center (http://brcweb.bio.cornell.edu).

Gel electrophoresis in fluorescent format was done on an automated DNA sequencer (Model 373A for AFLP analysis and Model 377 for *Hbr*-TD, Perkin-Elmer/ABI) as described by Casa et al. (2000). Electropherograms were analyzed using GeneScan v.2.1 (Perkin-Elmer/ABI). DNA fragments were sized by interpolation to an internal lane standard using the Local Southern algorithm. After verification of peak scoring for each DNA sample, GeneScan result files were imported into Genotyper v.2.5 (Perkin-Elmer/ABI). Polymorphic peaks between the parental lines were labeled manually, and a fragment presence/absence matrix was generated.

Linkage map

SSR markers were ordered by bin number according to the 1998 SSR Consensus Map (http://www.agron.missouri. edu). To this framework, *Hbr* and AFLP markers were integrated by the BC_2 algorithm and the 'distribute' command of Map Manager QTX b08 [http://mcbio.med.buffalo.edu/mmQTX.html; (Manly and Olson 1999)]. To include a locus in a linkage group, a minimum logarithm of odds (LOD) threshold of 3.0 and a maximum recombination fraction of 0.40 were used.

QTL analysis

To identify QTLs associated with grain yield, grain moisture or plant height, single-point analysis (SPA) and interval mapping (IM) were performed on genotypic data and phenotypic measurements collected at each location using QGENE software (Nelson 1997). For each trait, significance thresholds corresponding to a 1% or 5% type-I error for the whole genome were determined based on 5,000 permutations (Churchill and Doerge 1994).

To control local field heterogeneity when mapping QTLs, we also evaluated five spatial models of grain yield for each test site: response surface (RS), incomplete block (IB), autoregressive (AR), moving average (MA) and autoregressive moving average (ARMA). For IB analysis, we divided each trial site into blocks of 8 to 12 plots each. For RS analysis, we created orthogonal polynomial regression variables for rows and columns using ORPOL in SAS/IML (Wolfinger et al. 1997). Linear and quadratic row and column effects and their interactions (linear × linear, linear × quadratic, quadratic × linear, and quadratic × quadratic) were estimated by stepwise regression in which all non-significant ($p \geq 0.05$) variables were removed from the model. MA, AR and ARMA first-order, one-dimensional processes were applied in the directions of rows and columns (Binns 1987; Gleeson and Cullis 1987; Gleeson 1997).

Spatial analyses were executed using the STEPWISE (RS), GLM (RS, IB), AUTOREG (AR) and ARIMA (MA and ARMA) procedures in SAS v 8.1 (SAS Institute; Cary, N.C.). Mean-square (GLM and AUTOREG) and maximum-likelihood (ARIMA) estimates of genotypic, spatial, and error-variance components were obtained. For each trial site, the 159 markers were tested, one at a time, for the allele substitution effect on grain yield. All significant ($p < 0.05$) markers then were added back to each spatial model, and all-possible-subsets regression was performed using the RSQUARE procedure in SAS (SAS Institute; Cary, N.C.). We selected model size by adding another marker covariate until the R^2 improvement dropped below 25%. Augmented models were obtained by adding the set of significant markers with the largest R^2 for that number of covariates. QTL analysis and permutation tests then were performed on the residual of each augmented model using QGENE (Nelson 1997).

Because our objective was to improve RD6502 in hybrid combination with B73, we reported QTLs associated with favorable effects by RD3013 introgressions only: (1) grain yield QTLs detected at two or more locations using SPA or IM, (2) grain yield QTLs identified by spatial analysis but not by standard SPA or IM models without local environmental control, or (3) grain moisture and plant height QTLs detected by both SPA and IM. Estimates of QTL allele substitution effects (%) were calculated as 100-times the difference between the phenotypic mean of BC_2TC progenies with and without a given RD3013 introgression divided by the mean of BC_2TC progenies without the introgression. Because only one-half of the individuals in a BC_2TC plot derived from a heterozygous BC_2 individual would be heterozygous for a given QTL locus, a factor of two was included in the calculation (Xiao et al. 1998).

QTL-BC_3TC evaluation

For BC_3TC development, we targeted the three yield QTLs (*yld3.1*, *yld5.1* and *yld6.1*) that were identified in the BC_2TC at two or more locations using SPA ($p < 0.05$). In addition, we targeted a fourth yield QTL (*yld10.1*) that was detected only when a spatial component (IB, RS, AR, MA or ARMA) and three marker covariates were included in the regression model ($p < 0.05$). RD3013 alleles contributed positively to yield at all four of these target introgressions.

Fifty six BC_3S_1 individuals were selected based on four criteria: (1) each selection was homozygous for the RD3013 allele at one or more target introgressions, and its corresponding BC_2TC family (2) ranked among the top ten entries at two or more locations based on yield-to-moisture ratio, (3) outperformed the best check (B73 × RD6502 or Pioneer 3394) at two or more locations, and

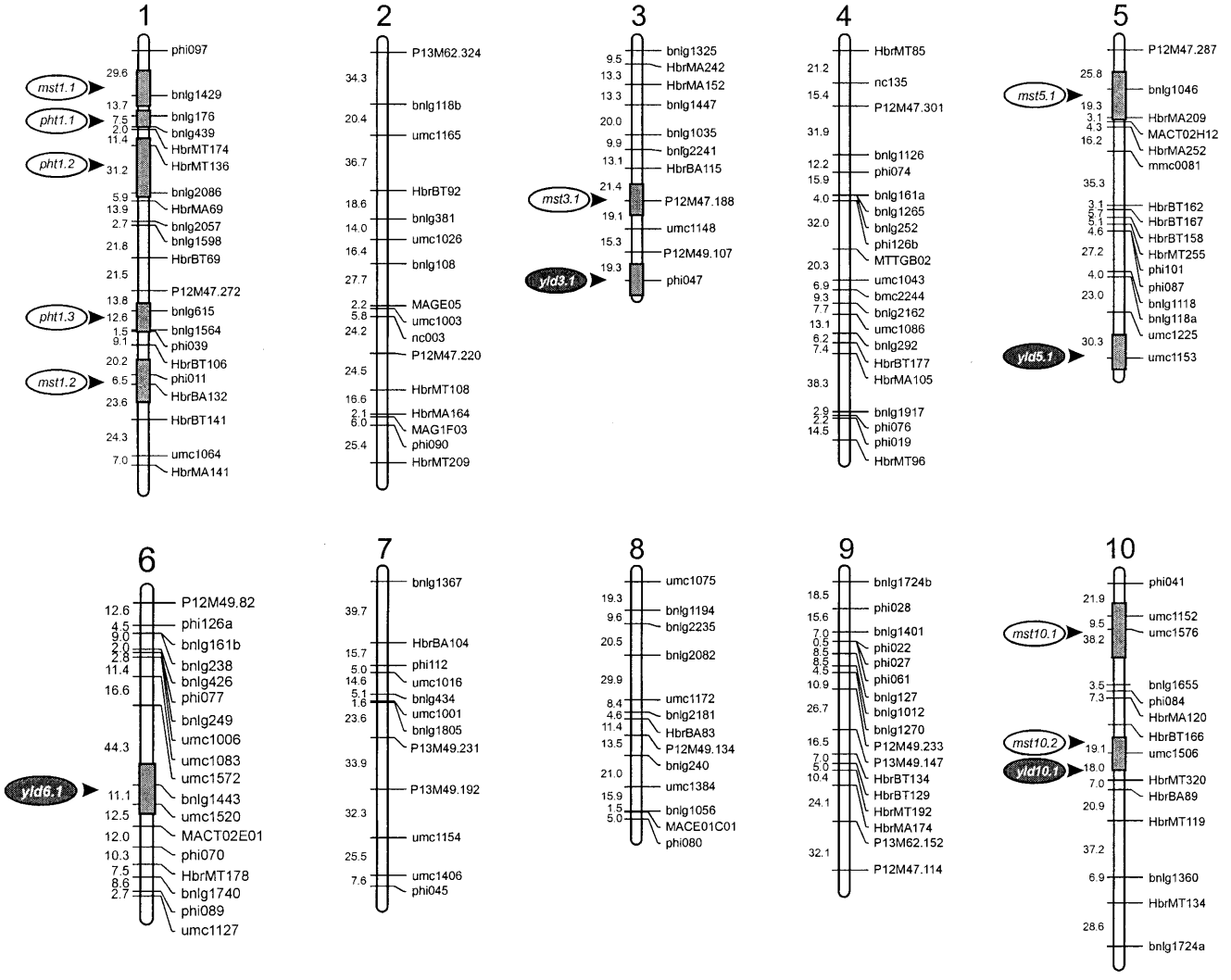

Fig. 1 Linkage map based on 204 BC_2 progeny derived from RD3013 × RD6502. Distances in cM (Kosambi) are to the left of each chromosome. AFLP and *Hbr* markers are named according to the primer combination, followed by the estimated size (bp) of the DNA fragment. *Boxes* along chromosomes encompass all markers associated with favorable effects by the RD3013 allele at a 5% empirical significance threshold: grain yield (*yld*) QTLs associated with positive effects by spatial analysis or by SPA at two or more locations and grain moisture (*mst*) or plant height (*pht*) QTLs associated with negative effects by SPA and IM. *Black-shaded ovals* indicate yield QTLs targeted for BC_3TC development

(4) showed acceptable standability where evaluated. Selected BC_3S_1 plants were backcrossed and crossed to B73 to generate 56 pairs of BC_4 and BC_3TC families, respectively.

In the year 2000, all BC_3TC families, the original B73 × RD6502 hybrid, and three local commercial checks were planted as two-row plots (5.3 × 0.76 m^2 at 69,200 plants/ha) with three replications at Pittsford, N.Y., Kingston, N.Y., Huxley, Iowa, and York, Neb., and with two replications at Jerseyville, Ill. A randomized complete block design (RCBD) was used in all locations with one plot per entry in each block. Weed and insect control and nutrient management recommendations for each local area were followed. Grain was combine- or hand-harvested (at Jerseyville only) and evaluated for grain yield and grain moisture.

In the year of this test, high winds and rain caused many plants to lodge prior to harvest, particularly in the Midwest. Therefore, we conducted diagnostic tests for autocorrelation and non-stationarity, and performed ARMA (1, 1) analysis on BC_3TC yield data using the ARIMA procedure in SAS v.8.1 (SAS Institute, Cary, N.Y.).

Results

Linkage map

We constructed a genetic map of 106 SSR, 15 AFLP, and 38 *Hbr* loci with an average marker distance of 14.4 cM (Fig. 1). The total length of the genetic map was 2,147.4 cM. Of the three marker systems, the *Hbr*-TD system was most efficient in detecting polymorphism between the parental lines. A total of 343 fragments were

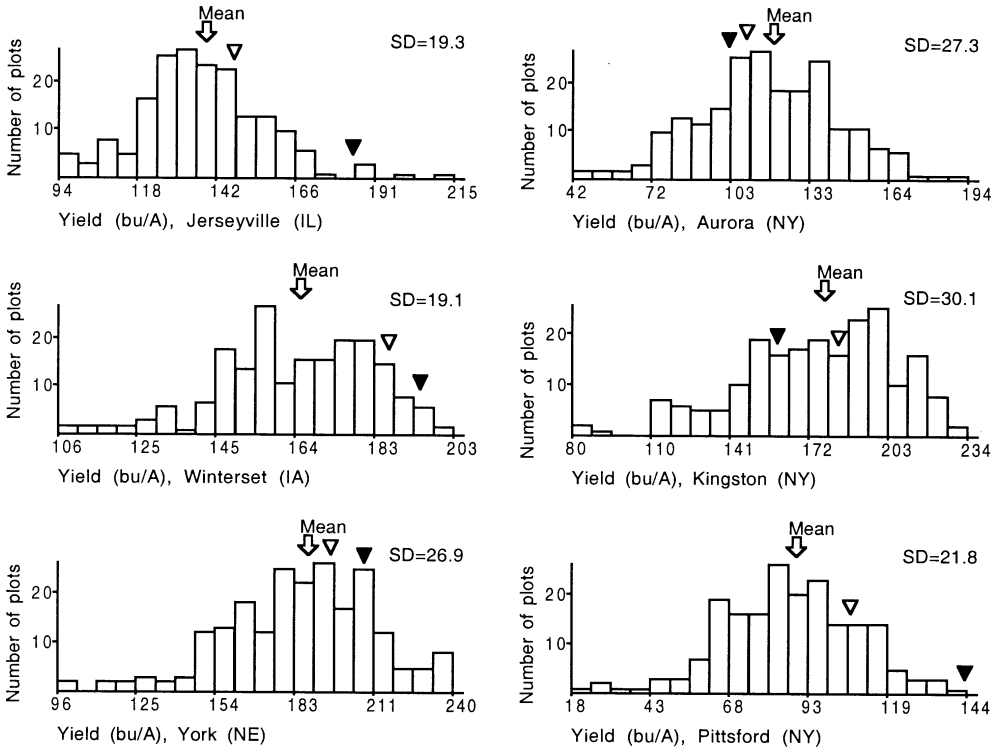

Fig. 2 Frequency histograms for grain yield of 204 BC_2TC progenies at six locations. *Arrows* indicate trial mean yield (adjusted to 15.5% grain moisture), triangles indicate means of the original hybrid B73 × RD6502 (*open triangle*) and Pioneer 3394 (*black-filled triangle*); SD = standard deviation

amplified using eight primer combinations (average of 43 bands per PCR reaction), 218 (64%) of which were unique to either RD6502 or RD3013. A lower level of polymorphism and a lower multiplex ratio were observed using the AFLP assay: 534 fragments were amplified by 14 primer combinations (average of 38 bands per PCR reaction), 206 (39%) of which were polymorphic. One hundred and twenty out of 204 (59%) SSR primer pairs detected polymorphism between the two parental lines.

Trait evaluation

Frequency histograms, means, and standard deviations for grain yield at the six locations are presented in Fig. 2. At all locations, we observed test hybrids that yielded higher than the original B73 × RD6502 hybrid and commercial check.

BC_2TC QTL analysis

Grain yield

We identified three QTLs (*yld3.1*, *yld5.1* and *yld6.1*) at which the RD3013 allele increased grain yield at two or more locations using SPA. A fourth QTL (*yld10.1*) was detected by spatial analysis but not by SPA or IM ($p < 0.05$; Table 1; Fig. 1). This RD3013 introgression also displayed a favorable, negative effect on grain moisture (*mst10.2*; Table 1; Fig. 1) at Pittsford and Kingston, N.Y. ($p < 0.05$).The estimated effect of each introgression on hybrid yield averaged 14.4% and ranged from 6.7% to 23.4% of the mean yield of non-carrier BC_2TC entries. R^2 values for the individual QTLs averaged 5.0%.

Spatial analysis confirmed the effects of *yld3.1* and *yld5.1*, but did not detect a significant effect on yield by RD3013 introgressions at *yld6.1* at any location. Except for *yld5.1* and *yld6.1* effects at Jerseyville, Ill, and *yld6.1* effects at York, Neb., spatial analysis resulted in *p*-values lower than that obtained by SPA.

Grain moisture

We identified six RD3013 alleles (*mst1.1*, *mst1.2*, *mst3.1*, *mst5.1*, *mst10.1* and *mst10.2*) associated with a desirable decrease in grain moisture by both SPA and IM ($p < 0.05$; Table 1; Fig. 1). The average decrease in moisture by individual QTLs was 7.2% of the average moisture of non-carrier BC_2TC progenies. R^2 values for individual QTLs averaged 6.2%.

Plant height

Three QTLs (*pht1.1*, *pht1.2*, *pht1.3*) at which the RD3013 allele displayed a desirable decrease in plant height were detected by SPA and IM ($p < 0.05$; Table 1; Fig. 1). The average decrease in plant height associated with individual QTLs was 5.8% of the average height of non-carrier BC_2TC progenies. R^2 values for individual QTLs averaged 5.8%.

Table 1 RD3013 alleles associated with favorable effects on yield, grain moisture, or plant height at a 5% empirical significance threshold

Trait	QTL	Bin(s)	SPA *p*-value	SPA LOD	% R^2	IM LOD	Spatial analysis *p*-value	Effect (%)[a]	Location
Grain yield[b]	*yld3.1*	3.09	0.0054**	1.70	4.0	1.70	0.0019**	–23.9	Aurora, N.Y.
			0.0390*	0.94	2.3	0.96	0.0327*	9.1	Winterset, Iowa
	yld5.1	5.09	0.0133*	1.34	3.3	1.83	0.0444*	12.1	Jerseyville, Ill.
			0.0194*	1.20	2.9	1.22	0.0126*	20.5	Pittsford, N.Y.
	yld6.1	6.05–6.06	0.0037**	1.86	4.8	2.28	0.1722	13.7	Jerseyville, Ill.
			0.0017**	2.17	5.3	2.42	0.5574	14.9	York, Neb.
	yld10.1	10.05	0.0788	0.68	12.1	0.88	0.0298*	6.7	Winterset, Iowa.
Grain moisture[c]	*mst1.1*	1.02	0.0023**	2.04	4.8	2.09*	n/a[d]	–5.9	Pittsford, N.Y.
	mst1.2	1.09–1.11	0.0007***	2.52	5.9	2.82*	n/a	–5.5	Aurora, N.Y.
	mst3.1	3.04–3.05	0.0012**	2.31	5.2	2.45*	n/a	–6.6	Aurora, N.Y.
	mst5.1	5.03–5.05	0.0014**	2.25	5.1	2.40*	n/a	–5.7	Kingston, N.Y.
	mst10.1	10.01–10.02	0.0001***	3.50	8.6	3.22*	n/a	–13.8	Winterset, Iowa.
	mst10.2	10.03–10.05	0.0004***	2.80	7.5	2.11*	n/a	–6.8	Pittsford, N.Y.
			0.0004***	2.71	6.1	2.81*	n/a	–5.9	Kingston, N.Y.
Plant height[c]	*pht1.1*	1.03	0.0005***	2.70	6.2	2.58*	n/a	–5.8	Kingston, N.Y.
	pht1.2	1.05–1.06	0.0007***	2.50	5.7	2.70*	n/a	–5.8	Kingston, N.Y.
	pht1.3	1.07	0.0010**	2.38	5.6	2.32*	n/a	–5.8	Kingston, N.Y.

*, **, *** $p<0.05$, 0.01, or 0.001

[a] Effect (%) = 200*(AB – AA)/AA; AB = mean of BC_2TC progenies with RD3013 allele; AA = mean of non-carrier BC_2TC progenies

[b] QTLs detected by SPA at two or more locations or by spatial analysis only

[c] QTLs detected by SPA and IM

[d] n/a = not evaluated

Table 2 RD3013 alleles associated with positive effects on BC_3TC yield in replicated field trials[a]

QTL	Marker	*Pr* > |*t*|	Estimated effect (% non-carrier)	Mean yield in bu/A		Location
				QTL-BC_3TC	Non-carrier	
yld3.1	*phi047*	0.0381*	11.1	168	140	Jerseyville, Ill.
yld10.1	*umc1506*	0.0228*	8.1	112	103	Huxley, Iowa.
		0.0490*	5.3	114	109	York, Neb.

*$p < 0.05$

[a] RCBD having one plot/entry/block × 56 entries × three blocks (except at Jerseyville where blocks = 2)

QTL-BC_3TC yield

For BC_3TC development, we targeted four RD3013 introgressions (*yld3.1*, *yld5.1*, *yld6.1* and *yld10.1*) associated with increases in BC_2TC yield. Diagnostic tests for autocorrelation and non-stationarity based on ACF, IACF, and PACF plots indicated correlation between yields of adjacent plots at all Midwestern locations; highly significant effects of AR or MA processes on hybrid yield were detected ($p < 0.01$).

ARMA analysis of BC_3TC yield in replicated trials confirmed the effects of two of the four target introgressions. Locations at which we confirmed *yld3.1* and *yld10.1* effects in the BC_3TC corresponded with the broad geographic region (i.e., Midwest versus N.Y.) in which they were identified in the BC_2TC. Specifically, BC_3TC entries containing *yld3.1* significantly outperformed non-carrier entries by 11.1% (15.6 bu/acre) at Jerseyville, Ill., ($p < 0.05$; Table 2) although realized gain was slightly less than that predicted (16.3%). BC_3TC entries containing *yld10.1* significantly outperformed non-carrier entries by an average of 6.7% (7.1 bu/acre) at Huxley, Iowa, and York, Neb.; in this case, realized gain matched the genetic gain (6.7%) predicted by BC_2TC analysis. Averaged across both target introgressions, the realized gain in yield was 84% the gain predicted by BC_2TC analysis.

High yielding test hybrids were identified based on BC_3TC field evaluation. For example, test hybrid B73 × 582-1 (161.6 bu/acre), containing *yld3.1*, outperformed the original hybrid B73 × RD6502 (137.0 bu/acre) by 18% at Jerseyville, Ill. Similarly, B73 × 537-4 (129.0 bu/acre) and B73 × 698-4 (128.7 bu/acre), each containing *yld10.*1, outperformed the original hybrid (95.8 bu/acre) by an average of 35% at Huxley, Iowa.

Contrary to expectation, *yld5.1* displayed a significant negative effect on yield at Pittsford (–7.8%; $p < 0.05$) where a positive effect had been detected in the BC_2TC (Table 1). In addition, *yld6.1* did not have a significant effect on BC_3TC yield at any location ($p \geq 0.05$).

Discussion

Identification of useful QTL alleles

The AB-QTL method is tailored to the discovery and transfer of favorable QTL alleles from unadapted germplasm to established lines (Tanksley and Nelson 1996). This work demonstrated that this breeding strategy can be extended to BC_2TC progeny derived from two elite, heterotic inbreds of maize for the identification and transfer of useful QTLs or novel epistatic gene interactions. For all three traits evaluated – grain yield, grain moisture and plant height – we identified QTL alleles in the Iodent-type donor parent RD3013 that improved hybrid performance when transferred to the Lancaster-type recipient parent RD6502. These results indicated favorable effects of RD3013 alleles in an RD6502 background or in a hybrid combination with B73 or both.

The validity of grain yield, grain moisture, and plant height QTLs identified in this experiment is supported by their co-localization with QTLs or structural genes mapped in other maize populations. Specifically, *yld6.1*, *mst5.1*, *pht1.1* and *pht1.2* QTLs mapped to the same chromosomal bins as previously reported yield (*qgyld8* and *qgyld17*), grain moisture (*qmoist3*), and plant height (*qpht61*, *qplht49*, *qplht56* and *qplht64*) QTLs, respectively (Pioneer Composite Map 1999; http://www.agron.missouri.edu). In addition, *pht1.1* mapped to the same chromosomal bin as *compact plant (ct2)*, a gene with known effects on plant stature in maize (Beavis et al. 1991). Although plant height typically shows quantitative inheritance, major genes for reduced plant height have been found in all major cereals (Börner et al. 1999), and QTLs associated with qualitative genetic loci controlling plant height have been reported (Beavis et al. 1991; Edwards et al. 1992; Berke and Rocheford 1995; Wu et al. 1996).

Methods of QTL analysis

Although we observed a good correspondence between results using SPA and IM methods of analysis, occasional discrepancies can reflect recombination between a marker locus and a QTL, the presence of a significant cofactor between two markers flanking any given QTL (e.g., cancellation of positive and negative dominance), a type-I error, or loss of statistical power upon adding another marker covariate to the regression model. In this study, we did not apply composite interval mapping to control residual genetic background. A BC_2 population is expected to have 87.5% of the recurrent parent genome recovered and, therefore, less variance due to genetic background compared to balanced populations. Moreover, the addition of each marker covariate to the regression model would cost a greater loss of QTL mapping power in an advanced backcross population due to the lower frequency of informative (i.e., recombinant) individuals for each marker locus.

Spatial analysis

As in most QTL mapping experiments, unreplicated field trials at multiple locations were conducted in order to maximize the number of BC_2 progeny that could be evaluated in this study. First, large sample sizes are required to detect genetic variance associated with many small-effect QTLs in the genome (Lande and Thompson 1990; Beavis 1994). Second, increasing sample size, more so than replication, improves the power of QTL mapping when significant genetic variance exists among lines nested in QTL genotypes (Knapp and Bridges 1990). Third, a single replication at many locations is superior to several replications at fewer locations in controlling genotype × environment (G × E) interactions (Weber 1980).

To control spatial heterogeneity among entries evaluated across a large field, we included a model selection procedure in which a fixed number of simple models of genotypic and spatial variance were searched for the lowest error variance and highest R^2. Highly significant effects of AR or MA processes on hybrid yield and greater power (e.g., no type-I error associated with *yld6.1*) and precision (e.g., lower *p*-values associated with *yld3.1*, *yld5.1* and *yld10.1* effects) when mapping QTLs in the BC_2TC clearly demonstrated the importance of local environmental control for QTL mapping in maize. Consistent *yld10.1* effects on BC_2 and BC_3 hybrid yield further illustrated the utility of spatial analysis for the discovery of genuine QTLs based on field data. This QTL was not detected using SPA or IM models without a local environmental component. Currently, we are conducting a more detailed assessment of the utility of combining QTL analysis and spatial statistics using data from maize field trials.

Efficiency of AB-QTL analysis in maize

The low R^2 values obtained in this study can be attributed to the presence of small-effect QTLs governing maize yield or low QTL mapping power in advanced backcross progeny. By Mendelian expectation, only 50% of the plants in each BC_2TC plot contain the donor parent allele for a given introgression. This "dilution effect" by non-carrier individuals can mask small effects of RD3013 QTL alleles occurring at low frequencies (approximately 25%) in the hybrids.

A more fundamental concern that must be investigated is whether genes with detectable and meaningful additive or dominant effects on hybrid yield are abundant enough in elite maize to justify the high cost of MAS. In this experiment, the effects of *yld3.1* and *yld10.1* on BC_3 hybrid yield averaged 8.2%. Given the 4 years required to implement the first cycle of AB-QTL breeding, this translates to an average annual gain of 2.1% compared to historical genetic gains estimated as 0.5 to 2% per year (Tiffany et al. 1992; Evans and Fischer 1999). In terms of annual gain in yield, the

AB-QTL method may prove cost-effective only when a small pyramid of two or three QTL alleles are transferred to each inbred parent of an elite hybrid.

On the other hand, the AB population structure offers three advantages over F_2 progeny of unrelated inbreds if high-order gene interactions play a significant role in the expression of key agronomic traits in maize. First, disruption of favorable epistatic interactions in an elite hybrid is minimized by initiating selection in an advanced backcross generation. Second, useful small-effect additive and dominant alleles can be identified which more likely would be confounded by epistasis in balanced populations. Third, simultaneous identification and transfer of QTLs by the AB-QTL method require less time than other strategies based on selfing for QTL detection and backcrossing for line development.

In replicated trials, we were able to validate favorable effects of two yield QTLs one generation after QTL discovery. By comparing yield means of BC_3TC carrier and non-carrier test hybrids, we can conclude that, while other genes may have been introgressed into RD6502, the gene(s) in the *yld3.1* and *yld10.1* chromosomal segments are primarily responsible for the observed differences in yield. These results support the underlying assumption of QTL breeding: beneficial QTL-alleles identified in segregating populations will continue to exert their positive effect in nearly isogenic backgrounds of the same recipient parent.

While the breeder's objective is to identify favorable QTL alleles with stable effects across environments, we evaluated mean values at each location rather than across the six locations. By this approach, we were able to detect both positive and negative effects of RD3013 introgressions; pooling data across environments could have obscured the effects of introgressions such as *yld5.1*, for which G × E interactions may have been significant. Unexpected effects of *yld5.1* can indicate loss of QTL alleles in the BC_3TC due to recombination, the type-I error when detecting QTLs in the BC_2TC, significant genotype × environment interactions (i.e., QTL × year), or epistasis. Climatic conditions were favorable for maize grain production in all BC_2TC test environments, but high winds caused severe lodging of plants in the evaluation of BC_3TC progeny. Expression of positive *yld5.1* effects on grain yield may have been dependent on favorable growing conditions (i.e., the BC_2TC trial) or confounded by poor standability in stress environments (i.e., the BC_3TC trial). By Mendelian expectation, the genetic background of BC_3TC progeny contained 6.25% more of the recurrent parent genome than the BC_2TC. Restoration of favorable epistatic interactions among RD6502 and B73 alleles or novel and negative epistatic interactions among RD6502, RD3013 and B73 alleles may have masked expected effects in the BC_3TC. Effects of opposite sign have been detected in other validation experiments (Melchinger et al. 1998).

Finally, we were able to select BC_3TC entries superior to the original hybrid and commercial check one generation after QTL identification. Parents of these high yielding test hybrids can provide a starting point for the fine mapping of genes underlying *yld3.1* and *yld10.1* effects. Estimating the effects of these chromosomal segments in different genetic backgrounds, and in combination with other Stiff Stalk Synthetic (SSS) testers, will clarify their utility for hybrid maize production. In summary, this work demonstrates that genetic gains by AB-QTL analysis can be coupled with the maintenance and selection of favorable epistatic gene interactions by traditional hybrid breeding for maize improvement.

Acknowledgements The authors thank M.J. Thomson, T. Fulton and J. Li for critical reading of the manuscript, and A. Casa for guidance in *Hbr*-TD. This work was supported in part by Monsanto, Mycogen, Dupont, a National Needs Graduate Fellowship in Plant Biotechnology (Federal agreement # 98-38420-5815), a DOE/NSF/USDA Collaborative Research in Plant Biology Grant (#DBI-9420687 Interdisciplinary Research Training Group in Molecular Mechanisms of Plant Processes), and a gift from Ms. Marie Lavallard.

References

Allard RW (1996) Genetic basis of the evolution of adaptedness in plants. Euphytica 92:1–11

Beavis WD (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. 49th Annual Corn and Sorghum Research Conf, pp 250–266

Beavis WD, Grant D, Albertsen M, Fincher R (1991) Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. Theor Appl Genet 83:141–145

Berke TG, Rocheford TR (1995) Quantitative trait loci for flowering, plant and ear height, and kernel traits in maize. Crop Sci 35:1542–1549

Binns MR (1987) Practical use of neighbor methods and alternatives in the analysis of field trials. Can J Plant Sci 67:477–489

Börner A, Korzun V, Malyshev S, Ivandic V, Graner A (1999) Molecular mapping of two dwarfing genes differing in their GA response on chromosome 2H of barley. Theor Appl Genet 99:670–675

Casa AM, Brouwer C, Nagel A, Wang L, Zhang Q, Kresovich S, Wessler SR (2000) The MITE family *Heartbreaker* (*Hbr*): molecular markers in maize. Proc Natl Acad Sci USA 97:10,083–10,089

Causse MA, Fulton TM, Cho YG, Ahn SN, Chunwongse J, Wu K, Xiao J, Yu Z, Ronald PC, Harrington SE (1994) Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138:1251–1274

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Darrah LL, Zuber MS (1986) United States farm maize germplasm base and commercial breeding strategies. Crop Sci 36:1109

Doebley J, Stec A, Gustus C (1995) *teosinte branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics 141:333–346

Edwards MD, Helentjaris T, Wright S, Stuber CW (1992) Molecular-marker-facilitated investigation of quantitative trait loci in maize. 4. Analysis based on genome saturation with isozyme and restriction fragment length polymorphism markers. Theor Appl Genet 83:765–774

Eshed Y, Zamir D (1994) Introgressions from *Lycopersicon pennellii* can improve the soluble-solids yield of tomato hybrids. Theor Appl Genet 88:891–897

Eshed Y, Zamir D (1996) Less-than-additive epistatic interactions of quantitative trait loci in tomato. Genetics 143:1807–1817

Eta-Ndu JT, Openshaw SJ (1999) Epistasis for grain yield in two F_2 populations of maize. Crop Sci 39:346–352

Eyre-Walker A, Gaut RL, Hilton H, Feldman DL, Gaut BS (1998) Investigation of the bottleneck leading to the domestication of maize. Proc Natl Acad Sci USA 95:4441–4446

Evans LT, Fischer RA (1999) Yield potential: its definition, measurement, and significance. Crop Sci 39:1544–1551

Gleeson AC (1997) Spatial analysis. In: Kempton RA, Fox PN (eds) Statistical methods for plant variety evaluation. Chapman and Hall, London

Gleeson AC, Cullis BR (1987) Residual maximum likelihood (REML) estimation of a neighbour model for field experiments. Biometrics 43:277–288

Knapp SJ, WC Bridges (1990) Using molecular markers to estimate quantitative trait locus parameters: power and genetic variances for unreplicated and replicated progeny. Genetics 126:769–777

Lamkey KR, Schnicker BJ, Melchinger AE (1995) Epistasis in an elite maize hybrid and choice of generation for inbred line development. Crop Sci 35:1272–1281

Lande R, Thompson R (1990) Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124:743–756

Laurie CC, True JR, Liu J, Mercer JM (1997) An introgression analysis of quantitative trait loci that contribute to the morphological difference between *Drosophila simulans* and *D. mauritiana*. Genetics 145:339–348

Long AD, Mullaney SL, Reid LA, Fry JD, Langley CH, Mackay TFC (1995) High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster*. Genetics 139:1273–1291

Lukens LN, Doebley J (1999) Epistatic and environmental interactions for quantitative trait loci involved in maize evolution. Genet Res 74:291–302

Manly KF, Olson JM (1999) Overview of QTL mapping software and introduction to Map Manager QT. Mammal Genome 10:327–334

Melchinger AE, Utz HF, Schön CC (1998) Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383–403

Moncada P, Martinez CP, Borrero J, Chatel M, Gauch H Jr, Guimaraes E, Tohme J, McCouch SR (2001) Quantitative trait loci for yield and yield components in an *Oryza sativa* × *O. rufipogon* BC_2F_2 population evaluated in an upland environment. Theor Appl Genet 102:41–52

Nelson JC (1997) QGENE: software for marker-based genomic analysis and breeding. Mol Breed 3:229–235

Panaud O, Chen X, McCouch SR (1996) Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). Mol Gen Genet 252:597–607

Stuber CW, Sisco PH (1991) Marker-facilitated transfer of QTL alleles between elite inbred lines and responses in hybrids. 46th Annual Corn and Sorghum Research Conf, pp 104–113

Tanksley SD, Nelson JC (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor Appl Genet 92:191–203

Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed Y, Petiard V, Lopez J, Beck-Bunn T (1996) Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. Theor Appl Genet 92:213–224

Tiffany GD, Smith OC, Salhuana W, Misevic D (1992) Considerations in selection of progeny from elite and broad-based germplasm. 47th Annual Corn and Sorghum Research Conf, pp 164–176

Vicente MC de, Tanksley SD (1993) QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134:585–596

Vuylsteke M, Mank R, Antonise R, Bastiaans E, Senior ML, Stuber CW, Melchinger AE, Lübberstedt T, Xia XC, Stam P, Zabeau M, Kuiper M (1999) Two high-density AFLP linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. Theor Appl Genet 99:921–935

Weber WE (1980) Quantitativ genetische Untersuchungsmethoden und ihre Bedeutung fur die Selektion in der Selbstbefruchterzüchtung. Habilitationsschrift, University of Hanover

Wolf DP, Hallauer AR (1997) Triple testcross analysis to detect epistasis in maize. Crop Sci 37:763–770

Wolfinger RD, Federer WT, Cordero-Brana O (1997) Recovering information in augmented designs, using SAS PROC GLM and PROC MIXED. Agron J 89:856–859

Wu P, Zhang G, Huang N (1996) Identification of QTLs controlling quantitative characters in rice using RFLP markers. Euphytica 89:349–354

Xiao J, Li J, Grandillo S, Ahn SN, Yuan L, Tanksley SD, McCouch SR (1998) Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. Genetics 150:899–909

Yu SB, Li JX, Xu CG, Tan YF, Gao YJ, Li XH, Zhang Q, Saghai Maroof MA (1997) Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. Proc Natl Acad Sci USA 94:9226–9231